

Ask Mr. Protocol

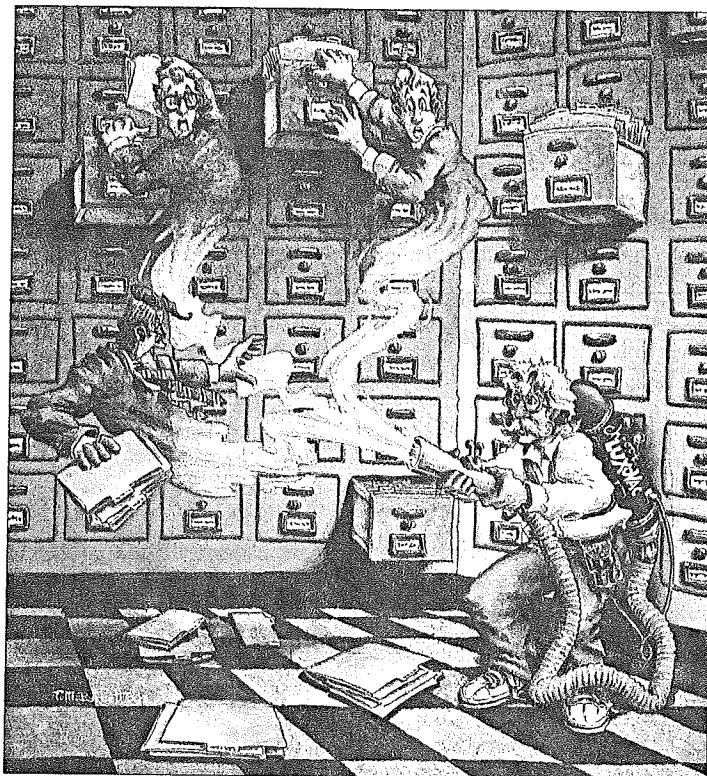


ILLUSTRATION BY TOM BARRETT

Wide-Area Services, or, Software By the Tankerload

by MICHAEL O'BRIEN

"I saved the world last Tuesday. It's your turn this week."

—A Network Services Provider

"Do you have an appointment?"

—A Non-Network Services Provider

"I've got a lorry-load of KCOPs waiting outside..."

—Tom Stoppard, *MacKoon's Hamlet*,
Kohout's Macbeth

Q I'm tired of grotting around trying to find out which anonymous FTP server has the software I want. How can I

figure out where to look? How do I find out which site has what? There must be a better way!

A: Yep, there must be. And when Mr. Protocol figures out what it is, he'll be the first to let you know about it. For now, though, anonymous FTP is about the best thing going. There are more helpful things on the horizon, though, and we'll be getting to them a little later on. These services are the first glimmers of what the network will look like when it's more thoroughly distributed, so they provide us with a bet-

ter glimpse of the future than most current services do. First, though, let's take a look at good old anonymous FTP. Mr. Protocol will, as usual, start with a historical perspective, which is not surprising, since he at least gives the impression of being older than most historical events himself.

In the beginning, when the net was so young that it wasn't even called the Internet, FTP was something that was only done between consenting adults. The model of the network, then as now, was that of mutually distrustful systems, so it was necessary to login to a remote system in order to use it.

OW! Um, Mr. Protocol has taken the opportunity to remind me that there were exceptions to this rule. A startling variety of innocent services have traditionally been available with-

out requiring login. Of these, just about the only surviving example is the "finger" protocol, but there have been many others. One very nice one was the "bank clock protocol," which provided the time and the temperature (inside and outside). This was handy for finding out whether the machine room air conditioning had in fact gone on vacation without having to truck on down to the machine room. There were many others, as Mr. Protocol has mentioned before. And some systems did not require a login at all—the old ITS system (which stood for Incompatible Timesharing System) took all comers. Every user had a distinct login, so that people could keep their files straight, but no password was required. In fact, the system would create new accounts automatically, on demand, for anyone who chose to come in over the network. This certainly made the machine a real network resource...too much of one for the powers that were, who demanded that a password scheme be implemented. Since they owned the road (literally), a minimal password scheme was instituted. ITS, however, to the end of its lifetime (which may not yet have arrived, since some Swedish students picked up the old MIT-MC hardware with the intention of continuing to run ITS on it) never had any file protection scheme.

We don't exactly digress here. ITS had no file protections, which meant that anyone could read, write or delete anything, whether a system file or a user file. ITS was regarded as the cooperative creation of everyone who used it. It also had the most non-standard user interface ever devised, bar none. Mr. Protocol once knew a fellow who was unable to figure out how to read a file on ITS by any other means than using FTP to send it to his terminal device. Being used to TENEX systems himself, he once hit Control-O to stop a file in mid-typeout. On ITS, Control-O meant "delete file," and that's just what ITS proceeded to do. Not the file he was looking at. Oh no. ITS deleted the file he was running. ITS deleted FTP. Frantically the fellow

paged an operator, who coolly replied, "Oh, yes, that happens all the time."

ITS had a superb, and often-used, file backup scheme.

FTP improved from there. From the beginning, there were operations that FTP could perform without requiring a login. Since in the early days, FTP carried the mail, this was one whole class of operations that could be carried out without a login. There were others, but transferring a file in either direction required you to have a login on the



machine. We now recognize the usefulness of not requiring this, but it was not so obvious in earlier times. Gradually, though, the usefulness of being able to pull files across the network from machines on which you did not have an account became evident. This occurred surprisingly late in the evolution of the Internet. Mr. Protocol feels that this is probably because in those days it was rare not to be able to get an account on any machine one desired, just by asking for it. Ah, youth.

The result is a peculiar institution in which every single machine on the Internet requests a name and a password, and in which in 98% of the cases, the answer is an evasion. It might be regarded as simpler to allow a file transfer from the anonymous area without needing a login at all, but because of the way FTP evolved, this isn't possible. In fact, it wasn't till FTP was converted to create a separate mode of operation that anonymous FTP became at all

widely supported. A user logging in anonymously (Mr. Protocol feels that this is an oxymoron if ever there were one) used to become a sort of ghost, free to wander about the foreign file system at will, but able to affect only those areas that were publicly writeable, and able to read only those areas that were publicly readable. This arrangement left too many security holes, so today's anonymous FTP was created. Most current FTP servers recognize the anonymous user as a special case, and proceed to use a special system call—called `chroot()` to deal with it. This system call changes the effective root of the file system to be the anonymous area. This acts as a firewall. Once this call has been made, the FTP server can no longer affect the rest of the file system, even if it becomes completely subverted. In effect, `chroot()` is a one-way trap door.

This makes anonymous FTP safe, but leaves the question of how to figure out where to get what. Those who inhabit the byways of the Internet have observed that there is a list of publicly available FTP sites, but even with this list, one has only the vaguest idea of who really carries what. One-line listings do not convey a large amount of information. There ought to be a way of automatically surveying all the anonymous FTP sites...and there is, or close to it.

This story is a bit complex. It starts with two fellows at McGill University, Alan Emtage and Peter Deutsch (no, this is NOT the L. Peter Deutsch of SDC/Xerox fame who now works for ParcPlace Systems), who very much agreed with the sentiments of the previous paragraph, and who therefore created an automatically updated database of every FTP site they could think of. That database now exists and is available at McGill. The next question is, how to get at it?

Mr. Protocol is glad you asked. The answer lies in one of those forward-looking services mentioned at the beginning of this screed: the Prospero file system.

The Prospero file system is the invention of B. Clifford Neumann of the University of Washington.

Prospero is defined as a computer protocol to permit the Internet to be viewed as a user-centered file system. Each user may configure an individual view of the Internet as a distributed file system, called a virtual file system. The contents of a directory may vary depending on the path by which that directory was reached.

The contents of a file or a directory may be modified by the addition of a filter, written in C, which modifies the behavior of the search or the apparent contents of the file or directory. All in all, the view of the distributed file system is completely customizable by the user. Prospero is implemented as a library that intercepts the usual `open()`, `close()`, `readdir()` and suchlike calls, and conditionally replaces them with calls to the Prospero virtual file system. An environment variable controls whether and under what conditions names are resolved in the native file system namespace or in the Prospero namespace.

And that is how the FTP database is made available. A Prospero client, named "archie," is freely available to interrogate the database, which is now usually referred to simply as the "Archie database." The archie program has the syntax of a text search utility. One merely hands it a keyword, or a more complicated query if desired, and archie uses the Prospero protocol to query the database at McGill. The result is a listing of each known host on the Internet whose anonymous FTP directory contains a file matching the query, together with the pathname of the file on that host (see "Dialoging with archie").

Certainly, archie could be implemented without the Prospero protocol. In fact, in the ordinary way of business, it would be implemented by applying for a TCP port number for the service, and defining a (much simpler!) protocol to send the query and the response. Possibly, like FTP, it would merely adopt the Telnet protocol. However, the implementors took the broader view that if the Prospero protocols were linked to a widely used application, experimentation with

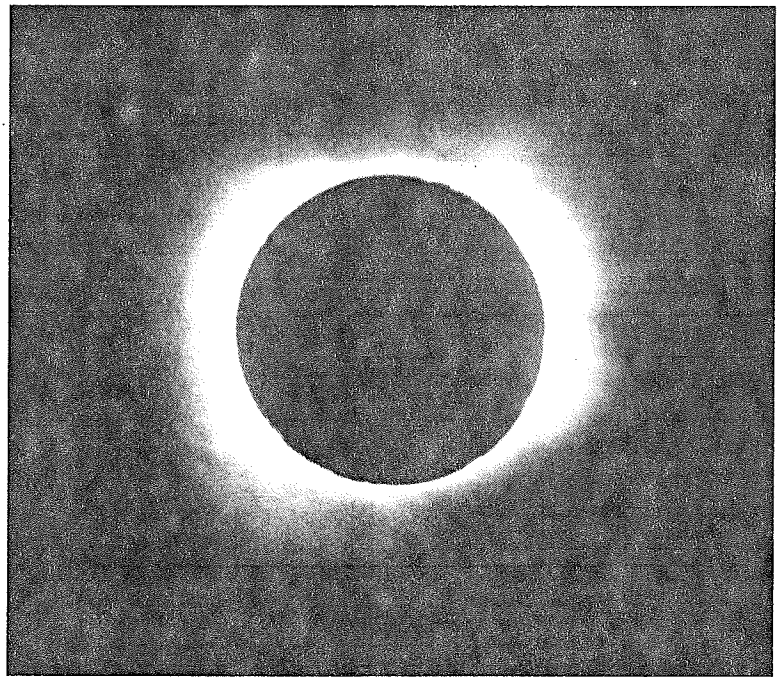
Prospero itself would be encouraged. Certainly the usefulness of the archie server proves that Prospero can do at least one thing well. The hope is that once a user puts the distribution together to run the archie client, other uses will readily become apparent. So far this seems to be working.

Prospero provides the user with a view of the Internet as a distributed file system...at least, that part of the Internet which has agreed to run Prospero and make material available.

This, together with the archie client software, serves to reduce some of the burden of locating and using information around the net. However, it does not solve the more general problem of locating arbitrary data sources around the network. This rather more formidable task has been undertaken by the Wide Area Information Servers (WAIS) project at Thinking Machines Inc., under Brewster Kahle.

Mr. Protocol, expressing a rare positive opinion of something new, thinks

What good is service that promises the moon when it means going for days without Sun.



When you consider how much you rely on your Sun system, it's good to know you can always replace it. Fast.

Apex Computer's 24-hour emergency program guarantees next-day replacement.

It's part of Apex's full menu of support services, including remote maintenance, depot repair and technical phone support. And it's all backed by warranty.

We'll do everything we can to keep you from going without Sun. Even if it takes a little moonlighting.



APEX
COMPUTER

We also buy & sell used Sun equipment.

Seattle • Sunnyvale • Boston • 1-800-654-8222

Dialoging with archie

```
% archie printf-scanf
```

```
021 Host speedy.cs.uiuc.edu
```

```
Location: /pub/MANCHESTER/july_update/usenet/printf-scanf.st
```

```
File -rw-r-r- 00013654 1991 Jun 20 15:04:00 GMT
```

```
printf-scanf.st
```

```
022 Host speedy.cs.uiuc.edu
```

```
Location: /pub/MANCHESTER/flat/printf-scanf.st
```

```
File -rw-r-r- 00013298 1990 Apr 30 00:00:00 GMT
```

```
printf-scanf.st
```

that the WAIS system is the first generally available glimpse of the Internet as it will someday appear. WAIS provides the user with a single generalized interface to information of almost any conceivable type, and as an added bonus, does this by use of an ISO protocol, rather than by requiring something so incompatible as to be automatically crossed off by standards-conscious managers everywhere. The protocol used is ANSI Z39.50-1988,

describing the Information Retrieval application-layer protocol. What WAIS actually does is provide a common interface allowing search and retrieval of data from all over the Internet.

The Information Retrieval protocol specifies how queries and responses may be sent, but it does not specify the actual usage of the application. WAIS turns out to be much more than just a database interface. In fact, if

WAIS were, say, a way of executing RPG II programs over the Internet, Mr. Protocol would doubtless be less amazed, and a great deal less interested. It seems as if people are always attempting to provide the most modern facilities for the most perverse and backward services. In fact, Mr. Protocol feels that many of the most widely advertised products in the "mainframe" computer press are the technical equivalent of laser-sighted, nuclear-powered siege engines.

This certainly does not include WAIS.

The central idea behind WAIS is similar to Prospero: treating the Internet as a sort of extended file system. In this case, though, the problem addressed is that of keeping information up-to-date. In the case of the archie server, it is the server's business to keep itself updated with respect to the contents of the various anonymous FTP archive sites around the net, which it does by interrogating each site periodically. It is the user's responsibility, however, to interrogate the archie server whenever updated information is wanted. WAIS takes

Expand the power of your LAN...



PPP – The wide area extension to your TCP/IP network.

PPP – Telecommute in style with demand dial, auto disconnect, packet filtering, inbound and outbound call support, and modem sharing with other applications.

Our business is customer support. It just so happens that we also sell high quality wide area communication solutions.

"I use Morning Star PPP to connect my business to the Internet. The package supports a large number of configurations, and is ideal for many IP connectivity applications. And best of all, Morning Star support is spectacular." — Marshall T. Rose

X.25/DDN – Additional wide area TCP/IP support is available with Morning Star's serial port and high-speed SnapLink X.25 products. Tomorrow's communication solutions are here today.

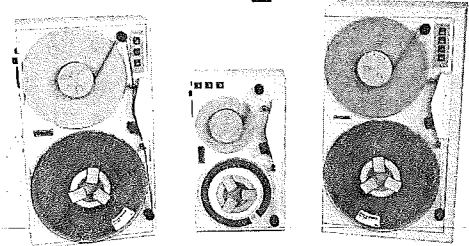
Morning Star Technologies

1760 Zollinger Road • Columbus, Ohio 43221

614 451 1883 • 800 558 7827 • Marketing@MorningStar.Com

Morning Star X.25, PPP and SnapLink™ are available on other UNIX™ platforms. SnapLink is a registered trademark of Morning Star Technologies, Inc. UNIX is a registered trademark of UNIX System Laboratories.

9-Track Tape Drives



QUALSTAR, the leading manufacturer of low-cost 9-track drives in the desktop marketplace, now offers these same cost-effective solutions for your workstation.

QUALSTAR provides complete 9-track tape subsystems ranging from 1600 to 6250 BPI in a compact package that's just right for your budget. Our solutions are available for SUN, AT&T, and MOTOROLA VME and SCSI based systems as well as the PC and MACINTOSH.

9-Track tape is the preferred choice for data interchange among professionals around the world. With a capacity of up to 250 megabytes, 9-track tape is also ideal for backup applications.

In just four years, QUALSTAR has become the dominant 9-track supplier in the desk-top marketplace. Find out why, call us today.

Dealer inquiries welcome!

QUALSTAR®

#1 Selling
9-Track Systems
on the Desktop



Qualstar Corporation
9621 Irondale Avenue
Chatsworth, CA 91311
FAX: 818 882-4081
Telephone: (818) 882-5822

©1989 BY QUALSTAR CORPORATION

All product and company names and trademarks are the exclusive property of their respective owners.

the view that collections of information should at least potentially be able to update themselves actively, so that the user can pose a query once, yet be able to see updated information each time he or she looks at the folder containing the query response.

WAIS queries are currently handled by keyword search, so questions can be phrased as English questions, or simply as lists of words having to do with the subject being queried. Additionally, when a document is discovered that the user finds to be a good match to his query, that document may be given as an example: "Find more documents like this one!"

What a user stores in the WAIS system then is not a view of a file system or a network, but a question, or a series of questions, which may be answered many times, in many ways. This facility of abstraction away from the details of network implementation is, Mr. Protocol feels, the best prediction yet of the future of network usage. More attention will be paid to the resulting information than to the protocol used to get it, at least by the user. Administrators, of course, will have to make sure that everything runs smoothly behind the scenes. Though he may not appear in public quite so regularly, it will be a long time before Mr. Protocol is out of a job. ➡

Mike O'Brien has been noodling around the UNIX world for far too long a time. He knows he started out with UNIX Research Version 5 (not System V, he hastens to point out), but forgets the year. He thinks it was around 1975 or so.

He founded and ran the first nationwide UNIX Users Group Software Distribution Center. He worked at Rand during the glory days of the Rand editor and the MH mail system, helped build CSNET (first at Rand and later at BBN Labs Inc.) and is now at an aerospace research corporation.

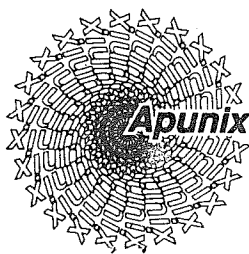
Mr. Protocol refuses to divulge his qualifications and may, in fact, have none whatsoever. His email address is amp@expert.com.

Present It Now!



Instant Full Color 35 mm Slides and Transparencies from Your Desktop

- Instant, full color 35 mm slides and small format overhead transparencies from your SPARCstation with the Polaroid C I 5000 Film printer.
- 4096 by 3600 resolution with 24 bits of color at each dot.
- Print files from any color PostScript application, such as Arts & Letters, with Apunix's NeWSprint driver.
- Sun Rasterfile filter for printing full color images in just minutes.
- SCSI interface attaches to Sun SCSI port for high speed transfer of image data.
- Turn all of your graphs, charts, and text material into eye-catching professional looking 35 mm slides and convenient pocket sized transparencies instantly, quickly, and inexpensively.



(800) 8AP-UNIX

Apunix Computer Services

**5575 Ruffin Road, Suite 110
San Diego, CA 92123**

Voice: (619) 495-9229 FAX: (619) 495-9230
UUCP: ...!ucbvax!ucsd!apunix!sales
Internet: apunix!sales@ucsd.edu

Our Family of products includes: Image Scanners • PostScript Printers • Color Printers
Serial I/O • Telebit Modems • Memory • Disks • Exabyte • DAT • Tape Stackers
Terminal Servers • Ethernet • QIC Tape • Half-inch Tape • Magneto Optical